DOCUMENT RESUME

ED 398 261 TM 025 215

AUTHOR Kitao, S. Kathleen; Kitao, Kenji
TITLE Testing Speaking.
PUB DATE [96]
NOTE 7p.
PUB TYPE Guides - Non-Classroom Use (055) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Communicative Competence (Languages); Educational Assessment; Foreign Countries; *Language Fluency; Language Proficiency; Language Tests; Language Usage; *Scoring; *Second Language Learning; *Speech Skills; *Test Construction; Test Items; Test Use

ABSTRACT

Speaking a second language is probably the most difficult skill to test in that it involves a combination of skills that may have no correlation with each other, and which do not lend themselves to objective testing. In addition, what can be understood is a function of the listener's background and ability as well as those of the speaker. Another difficulty is separating the listening skill from the speaking skill. In spite of the difficulties in testing speaking, it can be very beneficial in that it encourages the teaching of speaking in class. Reading aloud, conversational exchanges, and tests using visual material as stimuli are common test items for testing speaking. Oral interviews, role play tests, and group or pair activities are also useful. One of the great difficulties in testing speaking is the assessment and its scoring. If possible, the speaking tasks should be recorded and the scoring done from the tape. Aspects of speaking that might be considered in the assessment scale are grammar, pronunciation, fluency, content, organization, and vocabulary. Even though methods of testing speaking are not perfect, they are worth the effort for their effects on teaching and classroom instruction. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
S. KATHLEEN KITAO
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ED 398 261

Testing Speaking

S. Kathleen Kitao Doshisha Women's College Kyoto, Japan
Kenji Kitao Doshisha University Kyoto, Japan

Difficulties of Testing Speaking

Speaking is probably the most difficult skill to test. It involves a combination of skills that may have no correlation with each other, and which do not lend themselves well to objective testing. There are not yet good answers to questions about the criteria for testing these skills and the weighting of these factors. A speaker can produce all the right sounds but not make any sense, or have great difficulties with phonology and grammar and yet be able to get the message across. Also, success in speaking depends to a great extent on the listener. Comprehension of spoken material depends, among other factors, on the degree to which the listener is familiar with the speaker's accent and the degree to which they share background knowledge, and so what is a problem for one listener may not be a problem for another listener.

Another difficulty is separating the listening skill from the speaking skill. In most normal situations, there is an interchange between listening and speaking, and speaking appropriately depends, in part, on comprehending spoken input. This necessarily becomes a factor in the testing of speaking, and it is difficult to know whether you are testing purely speaking or speaking and listening together.

Testing speaking is also a particular problem when it is necessary to test large numbers of students. In some situations, it is necessary to test thousands of students, and even if each student speaks for only a few minutes, this becomes a huge job.

In spite of the difficulties inherent in testing speaking, a speaking test can be a source of beneficial backwash. If speaking is tested, unless it is tested at a very low level, such as reading aloud, this encourages the teaching of speaking in classes.

TM023215

Types of Speaking Tests

Reading Aloud

One way that reading is tested is by having the testees read aloud. This is not generally a good way to test reading. Its backwash effect is likely to be harmful, and it is not a skill that is used much outside of the classroom. However, it is a way to test pronunciation separately from the content of speech. If it is necessary to use this method of testing, the test should at least make use of a situation where the student might actually be reading aloud, such as reading instructions or parts of a letter to another person.

Conversational Exchanges

Another simple type of test is a test in which students are given a particular situation and instructed to respond in a certain way. These tests are usually highly structured and require only a limited response, not connected discourse.

Type 1. Students are asked to construct a sentence following a certain pattern using the information that they are given. They are usually given two or three examples first. For example:

1. Jane has been working at the same job for ten years. She is tired of her job and would like to do something more interesting. (*She wishes she could find a more interesting job.*)

2. It has been raining since yesterday. I want to go on a picnic this afternoon. (*I wish it would stop raining.*)

Type 2. In another type of conversational exchange test, a situation is given but no guidance is given as to how to respond, and the student can respond freely. For example:

A friend of yours has been tired a lot recently. What do you say to her? (*Maybe you should be getting more rest.*)

You have been served coffee in a restaurant but you haven't been given any cream or sugar, and you would like some. What do you say to the server? (*Could I have some cream and sugar.*)

Type 3. In another type of text, the student is given a stimulus sentence and can respond in any appropriate way. This type of test usually depends on conventional responses, such as responses to requests, invitations, etc.

Could I borrow your dictionary?
(*Sure, go ahead./ Sorry, I'm using it.*)

Hey, how are you doing?
(*Just fine, how are you?*)

This may also be done as part of a longer dialogue, where there are pauses for the testee's responses to questions on a tape. The problem with this sort of test is that the dialogue on the tape keeps going along, no matter what the testee says, whether the next line on the tape is appropriate or not. In addition, it is testing listening at least as much as it is testing speaking. However, this type of test does make it possible to test large numbers of people at the same time in a way that is at least somewhat communicative.

Testing Speaking Using Visual Material

Pictures, maps, diagrams, and other types of visual material can be used to test speaking without requiring the testee to comprehend written or spoken material. Through careful selection of the material, the tester can control the vocabulary and, to some extent, the grammatical structures required.

Various types of material are appropriate for this type of test, depending on the language that the tester wants to elicit. One common type of stimulus material is a series of pictures which tell a story, often with a few sentences of introduction to get the testee started. This requires the testee to put together a coherent narrative. A variation on this is to give the pictures in random order of the narrative to a group of testees. The students describe their pictures and decide on their sequence without showing them to each other, and then put them down in the order that they have decided on. They then have an opportunity to reorder the pictures if they feel it is necessary.

Another type of test using a visual stimulus is one in which two testees are given pictures that are similar but have several differences. Without seeing each other's pictures, they describe their own pictures and find the differences.

In addition to using pictures, it is possible to use charts, graphs, or other ways of visually organizing information, and the testee has to describe them. One possible problem with this type of visual stimuli is

that the material has to be something that all the testees can interpret equally well, since if a testee has difficulty understanding the visual information, that will influence the way he/she is evaluated.

Oral Interviews

Oral interviews are testing situations in which the tester and the testee carry on a conversation. The tester generally has a list of questions to ask the testee, and someone--either the interviewer or another person but preferably another person--assesses the language proficiency of the testee. The advantage of an interview is that it attempts to approximate a conversation situation, but the knowledge that it is a test and the constraints of the roles in the interview make it difficult to have a real, natural conversation.

The interview protocol needs to be carefully considered. It might be best to start with yes/no questions or factual questions that are easily answered to put the interviewee at ease, but these can be inhibiting, so this stage should not last too long. The questions should be interesting, easy to respond to, and give the testee some range in answering.

When doing an interview test, it is very important that the interviewer be well trained in order to avoid problems and so that the interviewer is consistent from one interview to another to help make the evaluation of the testee more reliable. The interviewer, for example, should avoid talking too much (though giving some personal information may help break the ice and put the testee at ease), should speak clearly, should not interrupt the interviewee to give answers, should maintain eye contact with the interviewee, and should not correct the interviewee during the interview.

Role Play Test

Another type of test is a role play. In a role play the testee and a confederate are given information on which to base a role play, and the testees are evaluated on their ability to carry out the task in the role play. For example, the role play might be getting information about course requirements. Role plays require the testee to use various functions that he/she might need in real communication.

Group or Pair Activities

To avoid the types of problems associated with oral interviews, sometimes students are put together in pairs or groups that are given some task, such as the ones mentioned above in the section on using visual materials. They might also be given a quiz to work on together, a puzzle

to work out, or a task. The point of these is not to find the "right" answer but to stimulate speech for the tester to evaluate. This is intended to be a less inhibiting situation than an interview test, but it also has the problem of lack of control. How the activity goes depends to a great extent on the mix of testees that are put together. If their level of English proficiency is widely different, or if one testee is particularly dominant, this will influence the interaction. A well-trained interviewer can provide similar environments from one interview to another, but this is not possible in testing a group of testees in a group activity. In addition, successful group discussions depend on factors other than speaking, including active listening and eliciting the opinions of all of the group members.

Assessment

One of the great difficulties in testing speaking is, of course, the assessment. It is necessary to develop a system of assessment that can be applied as objectively as possible, though it is probably never possible to avoid some subjectivity in assessment. The scale can be one general scale for overall speaking ability, or it can be divided between several aspects of the skill of speaking, such as pronunciation, grammar, organization, etc. The scale also depends on the speaking task that is used for the test. A test that uses public speaking as the task would be different from one that uses a group discussion.

If possible, the speaking task should be recorded and the scoring done from the tape. In addition, the marking should be done by more than one person and their reliability checked. If the task is an interviewer, the interviewer should not be required to score the test at the same time as conducting the interview, if this is avoidable.

Among the aspects of speaking that might be considered in the assessment scale are grammar, pronunciation, fluency, content, organization, content and vocabulary.

The band descriptions for a general scale might be as follows. The number indicates the level, and it is followed by a description of the characteristics of a speaker at that level.

7 Spoken communication is fluent, appropriate, and grammatically correct, with few if any errors.

6 Communication is generally fluent and grammatically correct with only occasional errors in grammar or pronunciation.

5 Student produces numerous grammatical errors and hesitations, but these do not interfere greatly with communication. Utterances are long and connected.

4 Student produces numerous grammatical errors and hesitations, and these occasionally interfere with communication. Utterances are short and connected.

3 Student's communication is limited to short utterances and depends in part on previously memorized conversational elements. Difficulty dealing with unpredictable elements. Many hesitations and grammatical errors. Communication only possible with sympathetic interlocutor.

2 Communication limited to short utterances, almost entirely memorized conversational elements. Unable to deal with unpredictable elements.

1 No communication possible.

Summary

Though speaking is a particularly difficult skill to assess, there are methods that can be employed to create situations that elicit speech and methods of assessing the testees' speech that are reasonably reliable. Testing speech is important for its backwash effect, even if the method of testing and of assessment are not as perfect as they might be.